



UNITED STATES PATENT AND TRADEMARK OFFICE

UNITED STATES DEPARTMENT OF COMMERCE
United States Patent and Trademark Office
Address: COMMISSIONER FOR PATENTS
P.O. Box 1450
Alexandria, Virginia 22313-1450
www.uspto.gov

APPLICATION NO.	FILING DATE	FIRST NAMED INVENTOR	ATTORNEY DOCKET NO.	CONFIRMATION NO.
09/587,990	06/06/2000	Chris A. Hamilton	024/1	8460

8791 7590 04/08/2005

BLAKELY SOKOLOFF TAYLOR & ZAFMAN
12400 WILSHIRE BOULEVARD
SEVENTH FLOOR
LOS ANGELES, CA 90025-1030

EXAMINER

ENG, GEORGE

ART UNIT PAPER NUMBER

2643

DATE MAILED: 04/08/2005

Please find below and/or attached an Office communication concerning this application or proceeding.

Office Action Summary	Application No.	Applicant(s)	
	09/587,990	HAMILTON, CHRIS A.	
	Examiner	Art Unit	
	George Eng	2643	

-- The MAILING DATE of this communication appears on the cover sheet with the correspondence address --

Period for Reply

A SHORTENED STATUTORY PERIOD FOR REPLY IS SET TO EXPIRE 3 MONTH(S) FROM THE MAILING DATE OF THIS COMMUNICATION.

- Extensions of time may be available under the provisions of 37 CFR 1.136(a). In no event, however, may a reply be timely filed after SIX (6) MONTHS from the mailing date of this communication.
- If the period for reply specified above is less than thirty (30) days, a reply within the statutory minimum of thirty (30) days will be considered timely.
- If NO period for reply is specified above, the maximum statutory period will apply and will expire SIX (6) MONTHS from the mailing date of this communication.
- Failure to reply within the set or extended period for reply will, by statute, cause the application to become ABANDONED (35 U.S.C. § 133). Any reply received by the Office later than three months after the mailing date of this communication, even if timely filed, may reduce any earned patent term adjustment. See 37 CFR 1.704(b).

Status

- 1) ☒ Responsive to communication(s) filed on 10 March 2005.
- 2a) ☐ This action is FINAL. 2b) ☒ This action is non-final.
- 3) ☐ Since this application is in condition for allowance except for formal matters, prosecution as to the merits is closed in accordance with the practice under *Ex parte Quayle*, 1935 C.D. 11, 453 O.G. 213.

Disposition of Claims

- 4) ☒ Claim(s) 5,6,9-14,29 and 30 is/are pending in the application.
- 4a) Of the above claim(s) _____ is/are withdrawn from consideration.
- 5) ☐ Claim(s) _____ is/are allowed.
- 6) ☒ Claim(s) 5,6,9-14,29 and 30 is/are rejected.
- 7) ☐ Claim(s) _____ is/are objected to.
- 8) ☐ Claim(s) _____ are subject to restriction and/or election requirement.

Application Papers

- 9) ☐ The specification is objected to by the Examiner.
- 10) ☐ The drawing(s) filed on _____ is/are: a) ☐ accepted or b) ☐ objected to by the Examiner.
Applicant may not request that any objection to the drawing(s) be held in abeyance. See 37 CFR 1.85(a).
Replacement drawing sheet(s) including the correction is required if the drawing(s) is objected to. See 37 CFR 1.121(d).
- 11) ☐ The oath or declaration is objected to by the Examiner. Note the attached Office Action or form PTO-152.

Priority under 35 U.S.C. § 119

- 12) ☐ Acknowledgment is made of a claim for foreign priority under 35 U.S.C. § 119(a)-(d) or (f).
- a) ☐ All b) ☐ Some * c) ☐ None of:
1. ☐ Certified copies of the priority documents have been received.
 2. ☐ Certified copies of the priority documents have been received in Application No. _____.
 3. ☐ Copies of the certified copies of the priority documents have been received in this National Stage application from the International Bureau (PCT Rule 17.2(a)).
- * See the attached detailed Office action for a list of the certified copies not received.

Attachment(s)

- | | |
|---|---|
| 1) <input checked="" type="checkbox"/> Notice of References Cited (PTO-892) | 4) <input type="checkbox"/> Interview Summary (PTO-413) |
| 2) <input type="checkbox"/> Notice of Draftsperson's Patent Drawing Review (PTO-948) | Paper No(s)/Mail Date. _____ |
| 3) <input type="checkbox"/> Information Disclosure Statement(s) (PTO-1449 or PTO/SB/08) | 5) <input type="checkbox"/> Notice of Informal Patent Application (PTO-152) |
| Paper No(s)/Mail Date _____ | 6) <input type="checkbox"/> Other: _____ |

DETAILED ACTION

Continued Examination Under 37 CFR 1.114

1. A request for continued examination under 37 CFR 1.114, including the fee set forth in 37 CFR 1.17(e), was filed in this application after final rejection. Since this application is eligible for continued examination under 37 CFR 1.114, and the fee set forth in 37 CFR 1.17(e) has been timely paid, the finality of the previous Office action has been withdrawn pursuant to 37 CFR 1.114. Applicant's submission filed on 3/10/2005 has been entered.

Response to Amendment

2. This Office action is in response to the amendment filed 3/10/2005. Accordingly, claims 1-4, 7-8, 15-28 and 31-34 are canceled and claims 5-6, 9-14 and 29-30 are pending for examination.

Claim Rejections - 35 USC § 103

3. The following is a quotation of 35 U.S.C. 103(a) which forms the basis for all obviousness rejections set forth in this Office action:

(a) A patent may not be obtained though the invention is not identically disclosed or described as set forth in section 102 of this title, if the differences between the subject matter sought to be patented and the prior art are such that the subject matter as a whole would have been obvious at the time the invention was made to a person having ordinary skill in the art to which said subject matter pertains. Patentability shall not be negated by the manner in which the invention was made.

Art Unit: 2643

4. Claims 5-6, 9, 11-14 and 29-30 are rejected under 35 U.S.C. 103(a) as being unpatentable over Terui et al. (US PAT. 5,684,527 hereinafter Terui) in view of Maekawa et al. (US PAT. 5,884,257 hereinafter Maekawa).

Regarding claim 5, Terui discloses a video conferencing system comprising a conference bridge (10, figure 1) for interconnecting a plurality of videoconference stations (12-18, figure 1) and a speaker identification subsystem (146, figure 7) to determine whether a conferee is speaking based on voice level and amount of motion (col. 3 lines 43-58, col. 5 lines 8-14 and col. 8 lines 1-4). In addition Terui also teaches the subsystem to compare level of voices made at each point in order to determine a speaker (col. 3 line 59 through col. 4 line 46 and col. 7 line 67 through col. 8 line 4) so that one skill in the art would recognize Terui teaching the subsystem to determine which of a plurality of conferees is speaking the loudest when multiple conferees are speaking simultaneously from different conference stations. Terui differs from the claimed invention in not specifically teaching the subsystem for determining whether a conferee is speaking based on whether lip movements of said conferee ascertained from a video signal from a conference station at which the conferee is located are reasonably consistent with an audio signal from the conference station. However, Maekawa teaches voice recognition apparatus for determining whether or not a speaker based on the voice signal and the lip movement signal, in order to prevent mis-recognition due to ambient noise (col. 5 line 11 through col. 7 line 37). Although Maekawa teaches to determine lip movement by radiating light from an LED onto a lip region instead of determining the lip movements ascertained from a video signal, Maekawa also teaches an alternative method for determining the lip movement by utilizing a camera for capturing an image of movement of lips, i.e., determining the lip movement ascertained from a

video signal (col. 12 line 60 through col. 13 line 6). Therefore, it would have been obvious to one having ordinary skill in the art at the time the invention was made to modify Terui in having the subsystem for determining whether a conferee is speaking based on whether lip movements of said conferee ascertained from a video signal from a conference station at which the conferee is located are reasonably consistent with an audio signal from the conference station, as per teaching of Maekawa, in order to prevent mis-recognition due to ambient noise.

Regarding claim 6, Terui disclose to compare voice levels at each point (col. 3 lines 61-63). Thus, the subsystem obviously comprises a voice activity detector.

Regarding claim 9, Maekawa teaches to including image analysis and recognition software (4, figure 1).

Regarding claim 11, Terui discloses a videoconference station (12, figure 1) comprising a transmitter (26, figure 1) to transmit a combined audio and video signal to a videoconference bridge (10, figure 1) and speaker identification subsystem (146, figure 7) located at the videoconference station to determine whether a conferee is speaking based on amount of motion (col. 3 lines 43-58, col. 5 lines 8-14 and col. 8 lines 1-4). In addition Terui also teaches the subsystem to compare level of voices made at each point in order to determine a speaker (col. 3 line 59 through col. 4 line 46 and col. 7 line 67 through col. 8 line 4) so that one skill in the art would recognize Terui teaching the subsystem to determine which of a plurality of conferees is speaking the loudest when multiple conferees are speaking simultaneously from different videoconference stations. Terui differs from the claimed invention in not specifically teaching the subsystem for determining whether a conferee located at the video conference station is speaking by analyzing whether lip movements of said conferee ascertained from a video signal at

said station are substantially consistent with an audio signal from the video conference station in which said conferee is located so as to produce human speech. However, Maekawa teaches voice recognition apparatus for determining whether or not a speaker based on the voice signal and the lip movement signal, in order to prevent mis-recognition due to ambient noise (col. 5 line 11 through col. 7 line 37). Although Maekawa teaches to determine lip movement by radiating light from an LED onto a lip region instead of determining the lip movements ascertained from a video signal, Maekawa also teaches an alternative method for determining the lip movement by utilizing a camera for capturing an image of movement of lips, i.e., determining the lip movement ascertained from a video signal (col. 12 line 60 through col. 13 line 6). Therefore, it would have been obvious to one having ordinary skill in the art at the time the invention was made to modify Terui in having the subsystem for determining whether a conferee located at the video conference station is speaking by analyzing whether lip movements of said conferee ascertained from a video signal at said station are substantially consistent with an audio signal from the video conference station in which said conferee is located so as to produce human speech, as per teaching of as per teaching of Maekawa, in order to prevent mis-recognition due to ambient noise.

Regarding claim 12, the limitations of the claim are rejected as the same reasons set forth in claim 6.

Regarding claim 13, the limitations of the claim are rejected as the same reasons set forth in claim 9.

Regarding claims 14, Terui discloses method of displaying images of a plurality of conferees (12-18, figure 1) in a videoconference system comprising the steps of determining

whether a conferee is speaking by analyzing audio signal and amount of motion from a conference station in which the conferee is located and visually altering an image of at least one of a plurality of remotely located conferees when said conferee is determined to be the loudest speaker (col.3 line 43 through col. col. 4 line 46 and col. 8 lines 1-4). In addition Terui also teaches the subsystem to compare level of voices made at each point in order to determine a speaker (col. 7 line 67 through col. 8 line 4) so that one skill in the art would recognize Terui teaching the subsystem to determine which of the conferees is speaking the loudest when multiple conferees are speaking simultaneously from different conference station. Terui differs from the claimed invention in not specifically teaching the step of determining whether a conferee located at the video conference station is speaking by analyzing whether lip movements of said conferee ascertained from a video signal from a conference station at which the conferee is located are substantially consistent with an audio signal from the video conference station in which said conferee is located so as to produce human speech. However, Maekawa teaches voice recognition apparatus for determining whether or not a speaker based on the voice signal and the lip movement signal, in order to prevent mis-recognition due to ambient noise (col. 5 line 11 through col. 7 line 37). Although Maekawa teaches to determine lip movement by radiating light from an LED onto a lip region instead of determining the lip movements ascertained from a video signal, Maekawa also teaches an alternative method for determining the lip movement by utilizing a camera for capturing an image of movement of lips, i.e., determining the lip movement ascertained from a video signal (col. 12 line 60 through col. 13 line 6). Therefore, it would have been obvious to one having ordinary skill in the art at the time the invention was made to modify Terui in having the step of determining whether a conferee located at the video

conference station is speaking by analyzing whether lip movements of said conferee ascertained from a video signal from a conference station at which the conferee is located are substantially consistent with an audio signal from the video conference station in which said conferee is located so as to produce human speech, as per teaching of Maekawa, in order to prevent mis-recognition due to ambient noise.

Regarding claim 29, Terui discloses the videoconferencing system further comprising means for visually altering an image of said conferee displayed in other conference station if said conferee is determined to be the loudest speaker of the plurality of conferees (col. 4 lines 22-46).

Regarding claim 30, the limitations of the claim are rejected as the same reasons set forth in claim 14.

5. Claim 10 is rejected under 35 U.S.C. 103(a) as being unpatentable over Terui et al. (US PAT. 5,684,527 hereinafter Terui) in view of Maekawa et al. (US PAT. 5,884,257 hereinafter Maekawa) as applied in claim 29 above, and further in view of Ogata et al. (JP 06062400A hereinafter Ogata).

Regarding claim 10, the combination of Terui and Maekawa differs from the claimed invention in not specifically teaching means for visually altering the image comprising means for highlighting a border around the image of the conferee determined to be the loudest speaker. However, Ogata teaches to display a red rectangular marker in a window display frame to indicate who is a speaker in order to easily specify who is a speaker (abstract). Therefore, it would have been obvious to a person of ordinary skill in the art at the time the invention was made to modify the combination of Terui and Maekawa in having means for highlighting a

Art Unit: 2643

border around the image of the conferee determined to be the loudest speaker, as per teaching of Ogata, in order to easily specify who is a speaker.

Response to Arguments

6. Applicant's arguments filed 3/10/2005 have been fully considered but they are not persuasive.

In response to applicant's argument that the combination of Terui and Maekawa fails to disclose the limitation of "a speaker identification subsystem to determine whether a conferee is speaking based, at least in part, on whether lip movements ascertained from a video signal from a conferee station ... are reasonably consistent with an audio signal from the conference station" due to Maekawa teach that lip movement is determined by radiating light from an LED onto a lip region such that any reflect light is detected by a photodiode, it is noted that Maekawa teaches the reason of using the combination of LED and the photodiode to detect lip movements because of cost effective, and Maekawa also teaches to use conventional method of utilizing a video camera for capturing an image of movement of lips to detect lip movements (col. 12 line 60 through col. 13 line 6) so that one skill in the art would recognizes Maekawa teaching to determine whether a person is speaking based on whether lip movements ascertained from a video signal are reasonably consistent with an audio signal. Thus, the amended claims are rendered obvious by the combination of Terui and Maekawa.

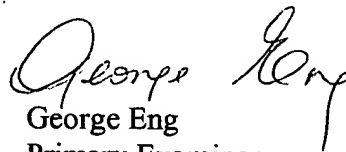
Conclusion

7. The prior art made of record and not relied upon is considered pertinent to applicant's disclosure. Stork et al. (US PAT. 5,771,306) disclose a speech recognition system utilizing dynamically varying acoustic and visual signals for improving the speech recognition system performance particularly in an adverse noisy environment (col. 3 line 61 through col. 4 line 62).

8. Any inquiry concerning this communication or earlier communications from the examiner should be directed to George Eng whose telephone number is (571) 272-7495. The examiner can normally be reached on Tue-Fri 7:30 AM-6:00 PM.

If attempts to reach the examiner by telephone are unsuccessful, the examiner's supervisor, Curtis A. Kuntz can be reached on (571) 272-7499. The fax phone number for the organization where this application or proceeding is assigned is 703-872-9306.

Information regarding the status of an application may be obtained from the Patent Application Information Retrieval (PAIR) system. Status information for published applications may be obtained from either Private PAIR or Public PAIR. Status information for unpublished applications is available through Private PAIR only. For more information about the PAIR system, see <http://pair-direct.uspto.gov>. Should you have questions on access to the Private PAIR system, contact the Electronic Business Center (EBC) at 866-217-9197 (toll-free).


George Eng
Primary Examiner
Art Unit 2643